AD-A231 023

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		Reprint

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Title shown on Reprint	DAAL03-89-K-0141
6. AUTHOR(S) Authors listed on Reprint	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Univ. of Wisconsin Madison, WI 53706	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Office P. O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSORING/MONITORING AGENCY REPORT NUMBER ARO 26460.6-EL

11. SUPPLEMENTARY NOTES

The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution unlimited.	

13. ABSTRACT (Maximum 200 words)

ABSTRACT SHOWN ON REPRINT

14. SUBJECT TERMS	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

ARO 26460.6-EL

Adaptive Detection in Subspaces

Barry D. Van Veen
Chong H. Lee

Department of Electrical and Computer Engineering
1415 Johnson Drive
University of Wisconsin
Madison, WI 53706

Abstract

This paper considers subspace based adaptive detection in the context of the likelihood ratio test studied by Kelly [1]. The probability of false alarm for this test depends only on the subspace dimension while the probability of detection is a function of the subspace. Thus, we propose choosing the transformation onto the subspace to maximize the probability of detection over a likely class of noise and interference scenarios. An approximate solution to this optimization problem is described. This approach can lead to dramatic increases in the probability of detection given a fixed number of data observations due to a large gain in the statistical stability associated with the reduced dimension subspace. The relationship between subspace design for adaptive detection and partially adaptive beamformer design is explored. Simulations verify the analysis.

1. Introduction

Processing of sensor array data in subspaces, sometimes termed the beamspace domain, has been proposed for adaptive beamforming and source location estimation (e.g. [2-4]). Benefits of mapping the data into a subspace prior to processing include reduced real time computational burden, improved adaptive convergence behavior, increased resolution and robustness. In this paper we consider subspace based adaptive detection for sensor array data in the context of the likelihood ratio test studied by Kelly [1]. The signal to be detected is modeled at the sensor outputs as a known vector with unknown amplitude. The noise and interference are modeled as zero mean Gaussian random vectors with unknown covariance matrix. The likelihood ratio test is dependent on the unknown signal amplitude and noise/interference covariance, so the detection statistic is obtained by maximizing the likelihood ratio over these unknown parameters given the data. Kelly [1] shows that the false alarm rate for this statistic is independent of the noise/interference covariance matrix and derives an expression for the probability of detection.

Kelly's [1] results are generalized here for detection in arbitrary reduced dimension subspaces. Expressions for the probabilities of detection and false alarm as a function of the subspace are given. The probability of false alarm is dependent only on the subspace dimension so we propose choosing the subspace to maximize the probability of detection over a likely class of interference scenarios. An approximate solution to this optimization problem is obtained by choosing the components of the subspace one at a time, with each additional component chosen to maximize the probability of detection. In spite of the loss in SNR which results from subspace processing, dramatic improvements in detection performance can be obtained because of the increased statistical stability associated with estimation of the covariance matrix in the reduced dimensioned subspace.

The outline of this paper is as follows. Section 2 introduces the detection problem and modifies Kelly's results for subspace processing. The relationship between partially adaptive beamforming and the parameters of the subspace detection problem is also illustrated. A technique for choosing the subspace to approximately maximize the probability of detection over a set of likely interference scenarios is described and analyzed in section 3. Several simulation examples are given in section 4 to illustrate the effectiveness of this approach. A summary is provided in Section 5.

2. Adaptive Detection

Let the N dimensional vector $\mathbf{x}_i$ be the data vector observed at the N sensor outputs at time i. The $\mathbf{x}_i$ are assumed to be independent and Gaussian distributed with means under the two hypotheses

$$H_o\!: E\{\mathbf{x}_i\} = 0,\ i = 0,1,\ldots,K$$
$$H_1\!: E\{\mathbf{x}_0\} = b\mathbf{s},\ E\{\mathbf{x}_i\} = 0,\ i = 1,2,\ldots,K \qquad (1)$$

The covariance is the same under both hypotheses

$$H_o, H_1\!: E\{(\mathbf{x}_i - E\{\mathbf{x}_i\})(\mathbf{x}_i - E\{\mathbf{x}_i\})^H\} = \mathbf{M} \qquad (2)$$

The signal vector $\mathbf{s}$ is assumed known, but the signal amplitude parameter b and the noise covariance $\mathbf{M}$ are unknown. Given K signal free data vectors $\mathbf{x}_i$, $i = 1,2,\ldots K$, we wish to determine whether or not a signal is present in the data vector $\mathbf{x}_0$.

Kelly [1] uses a likelihood ratio approach wherein the unknown parameters $\mathbf{M}$ and b are replaced by their maximum likelihood estimates. This leads to a threshold based decision statistic whose probability of false alarm (PFA) is dependent only on K, N and the threshold. An expression for the probability of detection is also derived in [1]. This approach is repeated below assuming the data is transformed into a subspace prior to performing detection.

Define an N by P (P< N) dimensioned full rank matrix $\mathbf{T}$ and transformed data vectors $\mathbf{z}_i = \mathbf{T}^H\mathbf{x}_i$, $i = 0,1,\ldots,K$. The subspace data remains zero mean except for $\mathbf{z}_0$ under H_1 which has mean $E\{\mathbf{z}_0\} = b\mathbf{T}^H\mathbf{s} = b\mathbf{s}_T$. The transformed covariance matrix is $\mathbf{M}_T = \mathbf{T}^H\mathbf{M}\mathbf{T}$. The joint probability density functions f_0 and f_1 under H_0 and H_1 are given by

$$f_0(\mathbf{z}_0,\ldots,\mathbf{z}_K) = \left\{\frac{1}{\pi^N \det\mathbf{M}_T}\exp\left[-\mathrm{tr}\left(\mathbf{M}_T^{-1}\mathbf{T}_o\right)\right]\right\}^{K+1} \qquad (3a)$$

$$f_1(\mathbf{z}_0,\ldots,\mathbf{z}_K) = \left\{\frac{1}{\pi^N \det\mathbf{M}_T}\exp\left[-\mathrm{tr}\left(\mathbf{M}_T^{-1}\mathbf{T}_1\right)\right]\right\}^{K+1} \qquad (3b)$$

where the sample covariance matrices $\mathbf{T}_0$ and $\mathbf{T}_1$ are

TH0331-9/90/0000-0163$01.00©1990 IEEE

$$T_0 = \frac{1}{K+1}\sum_{i=0}^{K} z_i z_i^H \tag{4a}$$

$$T_1 = \frac{1}{K+1}\left\{(z_0 - bs_T)(z_0 - bs_T)^H + \sum_{i=1}^{K} z_i z_i^H\right\} \tag{4b}$$

The likelihood ratio test is based on separate maximizations of f_0 and f_1 over the unknown parameters M_T and b. The detection statistic is then taken as the ratio of the maxima. The estimates of M_T and b which maximize f_0 and f_1 are by definition the maximum likelihood estimates.

It is well known that the maximum likelihood estimate of the covariance matrix is given by the sample covariance matrix [7]. Hence, T_0 maximizes f_0 over M_T and T_1 maximizes f_1 over M_T. We now have

$$\max_{M_T} f_0 = \left[(e\pi)^N \det T_0\right]^{K+1}$$

$$\max_{M_T} f_1 = \left[(e\pi)^N \det T_1\right]^{K+1} \tag{5}$$

b appears only in T_1 so for convenience we form the K+1[st] root of the ratio of f_1 to f_0, termed L(b), and maximize it over b to obtain the likelihood ratio test

$$L(b) = \frac{\det T_0}{\det T_1} \tag{6a}$$

$$\max_b L(b) = \frac{\det T_0}{\min_b \det T_1} \underset{H_0}{\overset{H_1}{\gtrless}} L_o \tag{6b}$$

where L_0 is the threshold parameter.

Using a derivation analogous to Kelly's [1], b is given by

$$b = \frac{s_T^H S^{-1} z_0}{s_T^H S^{-1} s_T} \tag{7}$$

where

$$S = \sum_{i=1}^{K} z_i z_i^H \tag{8}$$

Substitute for b in (6b) to obtain the likelihood ratio test

$$L = \frac{1 + z_0^H S^{-1} z_0}{1 + z_0^H S^{-1} z_0 - \frac{\left|s_T^H S^{-1} z_0\right|^2}{s_T^H S^{-1} s_T}} \underset{H_0}{\overset{H_1}{\gtrless}} L_o \tag{9}$$

Note that this test is of the same form as Kelly's, as it should be since the transformation T is linear. Thus, the analysis given by Kelly applies directly to (9) with minor modifications due to T. A summary of the properties of this test is given here; the reader is referred to [1] for details.

The probability of false alarm (PFA) is given by

$$PFA = \left(\frac{1}{L_0}\right)^{K+1-P} \tag{10}$$

This test has the constant false alarm rate (CFAR) property since the PFA depends only on K and P; it is independent of M and T. The probability of detection (PD) is

$$PD = 1 - \left(\frac{1}{L_0}\right)^L \sum_{k=1}^{L} \binom{L}{k} (L_0 - 1)^k H_k\left(\frac{a}{L_0}\right) \tag{11}$$

where L = K+1−P and

$$H_k(y) = \int_0^1 G_k(ry) f(r)\, dr \tag{12a}$$

$$G_k(x) = \exp(-x)\sum_{n=0}^{k-1} \frac{x^n}{n!} \tag{12b}$$

$$f(r) = \frac{(P+L-1)!}{L!(P-2)!}(1-r)^{P-2} r^L \tag{12c}$$

The parameter a which appears in the expression for PD is a function of the transformation matrix T

$$a = |b|^2 s_T^H M_T^{-1} s_T = |b|^2 s^H T\left(T^H M T\right)^{-1} T^H s \tag{13}$$

Transforming T by any nonsingular matrix on the right does not change a. The untransformed detection problem considered in [1] is represented by setting P = N.

The above expressions allow evaluation of the effect of T on detection performance. Mapping onto a subspace reduces the number of data vectors required for existence of the test. The original test required K > N; the transformed test requires K > P. If PFA is held constant, then reducing from N to P leads to a smaller threshold L_0 which tends to increase PD. r represents a Beta distributed (12c) loss factor which arises due to the estimation of the covariance matrix [1]. The loss factor reduces the effective value of a (see (12a)). Note that E{r} = (K+2−P)/(K+1). As P decreases, the distribution of the loss factor becomes concentrated closer to the maximum value of unity, which tends to increase PD. This increase corresponds to the gain in statistical stablity associated with estimation of the covariance matrix in the subspace. *Much fewer parameters* (P^2 vs. N^2) must be estimated given the same number of data vectors.

PD also increases as a increases. The dependence of a on T is illustrated by reexpressing (13) in the form

$$a = |b|^2 s^H M^{-H/2}\left[M^{H/2} T\left(T^H M^{1/2} M^{H/2} T\right)^{-1} T^H M^{1/2}\right] M^{-1/2} s \tag{14}$$

where $M^{1/2}M^{H/2}$ is the Cholesky factorization of M. The term in brackets is a projection matrix. Thus, a is proportional to the norm of the projection of $M^{-1/2}s$ onto the P dimensional space spanned by $M^{H/2}T$. This norm is upper bounded by T nonsingular (P = N). If M were known, we could attain the upper bound with P = 1 by choosing $T = M^{-1}s$. However, M is not known and we will not in general attain the upper bound. This results in a loss in performance associated with subspace processing.

The parameter a is related to the signal to noise ratio at the output of a linearly constrained minimum variance beamformer. Using the freedom to transform T on the right by a nonsingular matrix, we assume without loss of generality that T is of the form

$$T = [\, s \;\; T_s \,] \tag{15}$$

where the N by P−1 matrix T_s satisfies $s^H T_s = 0$. The partitioning in (15) is always possible provided $s^H T \neq 0$, a condition which is satisfied by all T of interest since a = 0 if $s^H T = 0$. Substituting this T into (13) yields

$$a = |b|^2 (s^H s)^2 \left(\begin{bmatrix} s^H M s & s^H M T_s \\ T_s^H M s & T_s^H M T_s \end{bmatrix}^{-1}\right)_{1,1} \tag{16}$$

where $(A)_{1,1}$ denotes the element in the first row and first column of $\mathbf{A}$. Apply the formula for the inverse of a partitioned matrix to obtain

$$a = \frac{|b|^2 (\mathbf{s}^H \mathbf{s})^2}{\mathbf{s}^H \mathbf{M s} - \mathbf{s}^H \mathbf{M T_s} (\mathbf{T_s^H M T_s})^{-1} \mathbf{T_s^H M s}} \quad (17)$$

a represents the signal to noise ratio at the output of a partially adaptive linearly constrained minimum variance beamformer with weight vector $\mathbf{w}$ expressed in generalized sidelobe canceller (GSC) [2] form as $\mathbf{w} = \mathbf{s} - \mathbf{T_s w_n}$. In GSC terminology, s is a fixed beamformer designed to pass the signal, $\mathbf{T_s}$ is the signal blocking matrix which is orthogonal to the signal, and $\mathbf{w_n}$ is a set of unconstrained adaptive weights. $\mathbf{w_n}$ is chosen to minimize the output power

$$\min_{\mathbf{w_n}} (\mathbf{s} - \mathbf{T_s w_n})^H E\{\mathbf{x x}^H\}(\mathbf{s} - \mathbf{T_s w_n}) \quad (18)$$

which has solution

$$\mathbf{w_n} = (\mathbf{T_s^H M T_s})^{-1} \mathbf{T_s^H M s} \quad (19)$$

This solution is obtained independently of whether $\mathbf{x}$ is distributed according to H_1 or H_0 because of the orthogonality of $\mathbf{T_s}$ and s. The output power due to the interference, P_I, is given by

$$P_I = \mathbf{s}^H \mathbf{M s} - \mathbf{s}^H \mathbf{M T_s} (\mathbf{T_s^H M T_s})^{-1} \mathbf{T_s^H M s} \quad (20)$$

and, assuming the signal is present in $\mathbf{x}$, the output power due to the signal, P_s, is

$$P_s = |b|^2 (\mathbf{s}^H \mathbf{s})^2 \quad (21)$$

Thus, (17) is rewritten

$$a = \frac{P_s}{P_I} \quad (22)$$

3. Design of T

The discussion in the previous section suggests a general philosophy for designing $\mathbf{T}$ to optimize the subspace based detector's performance. P should be chosen as small as possible and $\mathbf{T}$ should lead to values of a which are close to the upper bound. Below we propose two similar methods for designing $\mathbf{T}$.

Let θ be a vector which parameterizes the interference/noise environment and denote the corresponding covariance matrix as $\mathbf{M}(\theta)$. θ can represent interferer locations, power levels, isotropic background noise characteristics, etc. For example, if we assume the interference environment consists of two interferers in white noise, then θ would represent the interferer locations. Denote a as a function of θ as $a(\theta)$

$$a(\theta) = \mathbf{s}^H \mathbf{T} (\mathbf{T}^H \mathbf{M}(\theta) \mathbf{T})^{-1} \mathbf{T}^H \mathbf{s} \quad (23)$$

For convenience we assume $|b| = 1$. The goal is to optimize $\mathbf{T}$ over a set of likely interference scenarios denoted by Ω. For example, maximize a weighted average of $a(\theta)$ over Ω

$$\max_{\mathbf{T}} \int_{\theta \in \Omega} b(\theta)\, a(\theta)\, d\theta \quad (24)$$

where $b(\theta)$ is a nonnegative weighting function.

The correspondence between a and the SNR at the output of a partially adaptive beamformer can be used to pose an alternative optimization problem. $a(\theta)$ is inversely proportional to $P_I(\theta)$ (22) so maximizing $a(\theta)$ over $\mathbf{T}$ is equivalent to minimizing $P_I(\theta)$ over $\mathbf{T_s}$ where $\mathbf{T_s}$ is constrained to be orthogonal to the signal vector s. Here the constraint that $\mathbf{T}$ contain a component in the space spanned by s is explicitly enforced; this constraint is implicit in maximization of $a(\theta)$ over $\mathbf{T}$ although it can be made explicit by constraining $\mathbf{T}$ to be of the form given in (15). Minimizing a weighted average of the interference power over Ω gives the problem

$$\min_{\mathbf{T_s}} \int_{\theta \in \Omega} b(\theta)\, P_I(\theta)\, d\theta \quad (25)$$

This problem is identical to the partially adaptive beamformer design problem posed in [2]. Thus, the methods developed in [2,5,6] can be employed to solve (25).

Note that (24) and (25) are not equivalent. At a single value of θ, maximizing $a(\theta)$ over $\mathbf{T}$ is equivalent to minimizing $P_I(\theta)$ over $\mathbf{T_s}$. However, the optimization problems (24) and (25) perform the maximization and minimization over a range of θ. Using the relationship between $a(\theta)$ and $P_I(\theta)$ we see that (25) is equivalent to

$$\min_{\mathbf{T}} \int_{\theta \in \Omega} b(\theta) \frac{1}{a(\theta)}\, d\theta \quad (26)$$

Thus, while (24) maximizes the average SNR, (25) minimizes the average of the inverse SNR. The inverse SNR criterion more heavily weights small SNRs and deemphasizes large SNRs in choosing $\mathbf{T}$.

The remainder of this paper considers the optimization problem of (24). A closed form solution to (24) is not apparent so we follow a strategy similar to that posed for partially adaptive beamformer design [2]. Approximate solutions are obtained by designing $\mathbf{T}$ one column at a time. Partition $\mathbf{T}$ as $\mathbf{T} = [\,\mathbf{t}\ \ \mathbf{T_o}\,]$ where $\mathbf{t}$ is an N dimensional vector and $\mathbf{T_o}$ is assumed known. The P columns of $\mathbf{T}$ are designed one at a time, with each new column being a function of previously designed columns. $a(\theta)$ is expressed as a function of $\mathbf{t}$ and $\mathbf{T_o}$ using the relationship for the inverse of partitioned matrices as

$$a(\theta) = \mathbf{s}^H \mathbf{M}^{-1}(\theta)\mathbf{s} - \mathbf{c}^H(\theta)\mathbf{c}(\theta) + \frac{\mathbf{t}^H \mathbf{A}^H(\theta)\mathbf{c}(\theta)\mathbf{c}^H(\theta)\mathbf{A}(\theta)\mathbf{t}}{\mathbf{t}^H \mathbf{A}^H(\theta)\mathbf{A}(\theta)\mathbf{t}} \quad (27)$$

where

$$\mathbf{P}(\theta) = \mathbf{I} - \mathbf{M}^{H/2}(\theta)\mathbf{T_o}(\mathbf{T_o^H M}(\theta)\mathbf{T_o})^{-1}\mathbf{T_o^H M}^{1/2}(\theta) \quad (28a)$$

$$\mathbf{A}(\theta) = \mathbf{P}(\theta)\mathbf{M}^{H/2}(\theta); \quad \mathbf{c}(\theta) = \mathbf{P}(\theta)\mathbf{M}^{-1/2}(\theta)\mathbf{s} \quad (28b)$$

The first two terms in (27) do not depend on $\mathbf{t}$. Thus, the optimization problem is expressed as

$$\max_{\mathbf{t}} \int_{\theta \in \Omega_t} b(\theta) \frac{\mathbf{t}^H \mathbf{A}^H(\theta)\mathbf{c}(\theta)\mathbf{c}^H(\theta)\mathbf{A}(\theta)\mathbf{t}}{\mathbf{t}^H \mathbf{A}^H(\theta)\mathbf{A}(\theta)\mathbf{t}}\, d\theta \quad (29)$$

where Ω_t represents a subset of Ω.

Development of an approximate solution to an optimization problem of the same form as (29) is given in [2] and [5]. Application of these results leads to a $\mathbf{t}$ which satisfies the set of linear equations

$$\hat{\mathbf{A}}\mathbf{t} = \hat{\mathbf{c}} \quad (30)$$

where

$$\hat{\mathbf{A}} = \int_{\theta\in\Omega_t} b(\theta)\,\mathbf{A}^H(\theta)\mathbf{A}(\theta)\,d\theta \tag{31a}$$

$$\hat{\mathbf{c}} = \int_{\theta\in\Omega_t} b(\theta)\,\mathbf{A}^H(\theta)\mathbf{c}(\theta)\,d\theta \tag{31b}$$

Consider the case where the design region is approximated by a single point of the parameter space, $\Omega_t = \theta_o$ such that the integrals in (31a,b) are well approximated by the integrands evaluated at θ_o. Now $\hat{\mathbf{A}} = b(\theta_o)\mathbf{A}^H(\theta_o)\mathbf{A}(\theta_o)$ and $\hat{\mathbf{c}} = b(\theta_o)\mathbf{A}^H(\theta_o)\mathbf{c}(\theta_o)$. The presence of the projection matrix $\mathbf{P}(\cdot\,)$ in the definitions of $\mathbf{A}(\theta_o)$ and $\mathbf{c}(\theta_o)$ implies that $\mathbf{c}(\theta_o)$ lies in the space spanned by the columns of $\mathbf{A}(\theta_o)$ and that $\hat{\mathbf{A}}$ is not full rank. Let $\hat{\mathbf{A}} = \mathbf{U\Sigma V}^H$ be the singular value decomposition of $\hat{\mathbf{A}}$. The minimum norm solution to (30) is

$$\mathbf{t} = \mathbf{V}\Sigma^{-1}\mathbf{U}^H\mathbf{c}(\theta_o) \tag{32}$$

where here superscript -1 denotes pseudo inverse.

Compose $\bar{\mathbf{U}}$ of the columns of $\mathbf{U}$ corresponding to nonzero singular values. Using (32) in (27) we have

$$\begin{aligned} \mathbf{t}^H\mathbf{A}^H(\theta_o)\mathbf{A}(\theta_o)\mathbf{t} &= \mathbf{c}^H(\theta_o)\overline{\mathbf{U}\mathbf{U}}^H\mathbf{c}(\theta_o) \\ \mathbf{t}^H\mathbf{A}^H(\theta_o)\mathbf{c}(\theta_o) &= \mathbf{c}^H(\theta_o)\overline{\mathbf{U}\mathbf{U}}^H\mathbf{c}(\theta_o) \end{aligned} \tag{33}$$

so that

$$a(\theta_o) = \mathbf{s}^H\mathbf{M}^{-1}(\theta_o)\mathbf{s} - \mathbf{c}^H(\theta_o)\left(\mathbf{I} - \overline{\mathbf{U}\mathbf{U}}^H\right)\mathbf{c}(\theta_o). \tag{34}$$

Now $\bar{\mathbf{U}}$ is a basis for the space spanned by the columns of $\mathbf{A}(\theta_o)$. $\mathbf{c}(\theta_o)$ lies in this space so the projection of $\mathbf{c}(\theta_o)$ onto the space orthogonal to $\bar{\mathbf{U}}$ is zero and (34) simplifies to

$$a(\theta_o) = \mathbf{s}^H\mathbf{M}^{-1}(\theta_o)\mathbf{s}. \tag{35}$$

which is the maximum value for $a(\theta_o)$.

Thus, as the design region shrinks to a point, the design procedure suggested here leads to a value for $a(\theta_o)$ which is equal to the value obtained in the absence of subspace processing. This suggests that the best overall performance is obtained by designing each column of $\mathbf{T}$ over separate subregions of Ω which are are made as small as possible. If these subregions are sufficiently small, then the subspace based detector should show very little degradation in $a(\theta)$.

The correspondence between a and the SNR at the output of a partially adaptive beamformer facilitates analysis of the sensitivity of the design procedure to differences between the actual interference environment and the class of environments assumed during design. This allows determination of which parameters are most important to include in θ. If the design results in $a(\theta)$ being large for θ in the set Ω, then $P_I(\theta)$ will be small for θ in the set Ω. We address the sensitivity of $a(\theta)$ to environments θ not in Ω by examining the sensitivity of $P_I(\theta)$ to environments θ is not in Ω.

Suppose the interference environment consists of J point interferers in white noise so that

$$\mathbf{M}(\theta) = \sum_{i=1}^{L} \sigma_i^2\mathbf{d}(\alpha_i)\mathbf{d}^H(\alpha_i) + \sigma_w^2\mathbf{I} \tag{36}$$

where $\mathbf{d}(\alpha)$ is the array response or direction vector corresponding to direction α. In general, θ for this environment would include L, σ_w^2, σ_i^2, and α_i, $i = 1, 2, \ldots, L$. $P_I(\theta)$ is expressed as

$$\begin{aligned} P_I(\theta) &= \left(\mathbf{s} - \mathbf{T}_s\mathbf{w}_n\right)^H\mathbf{M}(\theta)\left(\mathbf{s} - \mathbf{T}_s\mathbf{w}_n\right) \\ &= \sum_{i=1}^{L} \sigma_i^2\left|\mathbf{s}^H\mathbf{d}(\alpha_i) - \mathbf{w}_n^H\mathbf{T}_s^H\mathbf{d}(\alpha_i)\right|^2 + \sigma_w^2\left|\mathbf{s} - \mathbf{T}_s\mathbf{w}_n\right|^2 \end{aligned} \tag{37}$$

Note that $\mathbf{w}^H\mathbf{d}(\alpha)$ is the response in direction α associated with $\mathbf{w}$.

There are two factors in P_I: the interference term depend on how well the response of s matches the response of $\mathbf{T}_s\mathbf{w}_n$ at the interferer directions α_i, and the white noise term depends on the norm of $\mathbf{s} - \mathbf{T}_s\mathbf{w}_n$. In general these two terms are inversely related; as response matching improves the norm gets larger. The relative sizes of σ_i^2 and σ_w^2 determine the relative importance of response matching versus weight vector norm.

If the σ_i^2 are much larger than σ_w^2, then response matching dominates. Given a $\mathbf{T}_s$ which provides good response matching, P_I will be insensitive to variations in σ_i^2. P_I will be sensitive to variations in the interferer directions. If $\mathbf{T}_s$ is chosen to provide good response matching over one range of α, it is unlikely that good response matching will be obtained over a different range of α. P_I should not be sensitive to changes in the number of interferers L provided $L \le P-1$. Assuming the $\mathbf{d}(\alpha_i)$, $i = 1,2,\ldots,L$ are linearly independent, the response matching term represents L equations in P–1 unknowns ($\mathbf{w}_n$). If $L \le P-1$, then the response matching term can always be made equal to zero independent of $\mathbf{T}_s$, although this may result in a large norm for $\mathbf{s} - \mathbf{T}_s\mathbf{w}_n$. A good $\mathbf{T}_s$ will achieve response matching close to zero while keeping the norm small.

If σ_w^2 is much larger than the σ_i^2, then the norm dominates and the design will not be sensitive to variations in σ_i^2, to L, or the interferer locations. A $\mathbf{T}_s$ designed for σ_i^2 much larger than σ_w^2 should perform well when σ_w^2 is much larger than the σ_i^2 because a $\mathbf{T}_s$ which results in good response matching can yield a small norm by setting $\mathbf{w}_n = \mathbf{0}$. In contrast, a $\mathbf{T}_s$ designed for σ_w^2 much larger than the σ_i^2 is not necessarily capable of good response matching.

The preceding discussion suggests that the most important parameters to include in θ are the interferer directions. The best all around performance is obtained by designing $\mathbf{T}_s$ assuming the σ_i^2 is much larger than σ_w^2. The exact number of interferers does not appear to be significant as long as $L \le P-1$.

4. Simulations

A linear equal spaced array of 50 sensors is used to illustrate the potential performance improvements which are possible by performing detection in subspaces. The sensors are spaced at one half wavelength. The signal to be detected arrives from the direction perpendicular to the array so

$$\mathbf{s} = 50^{-1/2}[\,1\ 1\ 1\ \ldots\ 1]^H.$$

The magnitude of b is set to unity. The noise consists of two interferers in white noise, where the ratio of interferer power to white noise power is 30 dB. This implies

$$M = 1000d(\alpha_1)d^H(\alpha_1) + 1000d(\alpha_2)d^H(\alpha_2) + 50^{-1/2}I \quad (38)$$

where $d(\alpha)$ is assumed to have unit norm and α_1, α_2 represent the interferer directions.

The subspace based detector is designed for $P = 4$ by solving (30) for each of the columns of T. $b(\theta)$ is equal to unity and the integrals in (31a,b) are approximated by sums. The elements of θ are α_1 and α_2 and T is designed to optimize performance over the range $.1 \le \sin\alpha_1 \le 1$, $.1 \le \sin\alpha_2 \le 1$. The subregion over which each column of T is designed is given in Table 1. These regions were not customized to optimize performance, but were selected in a somewhat arbitrary manner.

The values of a for the subspace and non-subspace detectors are computed at 182 different combinations of α_1 and α_2 on the region $.1 \le \sin\alpha_1 \le 1$, $.1 \le \sin\alpha_2 \le 1$. Table 2 presents a histogram of the ratio of non-subspace a to subspace a measured in dB. This number is a measure of the loss in a resulting from subspace processing. The design procedure is clearly effective since 177 of the 182 cases show less than 1 dB loss due to the mapping from a 50 dimensional space to a 4 dimensional space. The worst case loss is 3.51 dB.

The PD's of both detectors are computed for the same 182 cases assuming K = 75 signal free data vectors and a constant PFA = 10^{-6}. The PD's for the non-subspace detector ranged from a minimum of .4855 to a maximum of .5066. Table 2 presents a histogram of the PD's for the subspace detector. Subspace processing results in a tremendous improvement in the PD; the PD is greater than .999 for 177 of the 182 cases. The worst case PD of .8539 is significantly greater than the best case non-subspace PD of .5066. The two cases where the PD is less than .9 correspond to the two cases where the loss in a is greater than 1.5 dB. Customization of the regions used to design each column of T could be used to improve a and the PD at these points.

A histogram of the PD's over the same set of interference scenarios is given in Table 4 assuming PFA = 10^{-6} and K = 30. The non-subspace detector does not exist in this case because K < N. The worst case PD is .6472; 175 of the cases have PD's better than .99. The subspace detector performs significantly better than the non-subspace detector even though it has less than one half the number of signal free data vectors available.

5. Summary

The detection problem posed by Kelly [1] is modified to perform detection on data which is mapped into a subspace prior to processing. The detection performance tends to increase due to the reduction in data dimension, but tends to decrease due to a loss in SNR associated with the mapping into the subspace. A procedure is proposed for designing the subspace transformation to minimize the SNR loss. The subspace design problem for optimizing detection is shown to be closely related to the partially adaptive beamformer design problem. Simulations illustrate the effectiveness of subspace detection. The gain in detection performance associated with reducing dimension far exceeds the detection loss associated with the loss of SNR in the subspace for the cases studied here.

References

[1] E.J. Kelly, "An adaptive detection algorithm," *IEEE Trans. AES*, vol. AES-22, March 1986.

[2] B.D. Van Veen and R.A. Roberts, "Partially adaptive beamformer design via output power minimization," *IEEE Trans. ASSP*, vol. ASSP-35, pp. 1524-1532, Nov. 1987.

[3] P. Forster and G. Vezzosi, "Application of spheroidal sequences to array processing," *ICASSP 87*, pp. 2268-2271, April 1987.

[4] H.B. Lee and M.S. Wengrovitz, "Resolution threshold of beamspace MUSIC for two closely spaced emitters," to appear *IEEE Trans. ASSP*.

[5] B.D. Van Veen, "An analysis of several partially adaptive beamformer designs," *IEEE Trans. ASSP*, vol ASSP-37, pp. 192-203, Feb. 1989.

[6] B.D. Van Veen, "Optimization of quiescent response in partially adaptive beamformers," *IEEE Trans. ASSP*, vol. ASSP-38, pp. 471-477, March 1990.

[7] N.R. Goodman, "Statistical analysis based on a certain multivariate Gaussian distribution," *Ann. Math. Stat.*, vol. 34, pp. 152-177, March 1963.

Table 1. Design regions for the columns of T.

Column	$\sin\alpha_1$	$\sin\alpha_2$
1	.6 - 1.0	.6 - 1.0
2	.6 - 1.0	.1 - .58
3	.3 - .58	.1 - .58
4	.1 - .28	.1 - .28

Table 2. Difference between 10log(a) for the non-subspace detector and 10log(a) for the subspace detector.

Range for Difference	Number of Cases
0 < dif < .5	174
.5 < dif < 1.0	3
1.0 < dif < 1.5	3
1.5 < dif < 3.6	2

Table 3. PD's for subspace detector with K = 75. The non-subspace detector PD's range from .4855 to .5066.

Range for PD	Number of Cases
.85 < PD < .9	2
.9 < PD < .99	0
.99 < PD < .999	3
.999 < PD < .9999	4
.9999 < PD < .99999	173

Table 4. PD's for subspace detector with K = 30.

Range for PD	Number of Cases
.64 < PD < .9	2
.9 < PD < .99	4
.99 < PD < .999	175